PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 9/00 | A2

(11) International Publication Number: **WO 99/02659**

(43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/EP98/04265

(22) International Filing Date: 9 July 1998 (09.07.98)

(30) Priority Data:
9714475.2 9 July 1997 (09.07.97) GB

(71) Applicant *(for all designated States except US)*: BIOCHEMIE GESELLSCHAFT M.B.H. [AT/AT]; A–6250 Kundl (AT).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: SCHÖRGENDORFER, Kurt [AT/AT]; Niederbreitenbach 172, A–6322 Langkampfen (AT). STEMPFER, Günther [DE/AT]; Karl Schönherrstrasse 9/37, A–6330 Kufstein (AT). HUSI, Holger [CH/GB]; 13 Sciennes Road, Edinburgh EH9 1LG (GB). WALKINSHAW, Malcolm, Douglas [GB/GB]; 68 Spylaw Bank Road, Edinburgh EH13 0JB (GB).

(74) Agent: BECKER, Konrad; Novartis AG, Patent– und Markenabteilung, Lichtstrasse 35, CH–4002 Basel (CH).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: METHOD OF ALTERING THE DOMAINS OF CYCLOSPORIN SYNTHETASE TO GIVE A MODIFIED CYCLOSPORIN SYNTHETASE

(57) Abstract

The method of altering the domains of cyclosporin synthetase, to modified cyclosproin synthetase, to methods for the production of cyclosporin–like peptides or derivatives using the modified cyclosporin synthetase and to the cyclosporin–like peptides or derivatives.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

METHOD OF ALTERING THE DOMAINS OF CYCLOSPORIN SYNTHETASE TO GIVE A MODIFIED CYCLOSPORIN SYNTHETASE

This invention relates to a method of altering the domains of cyclosporin synthetase, to modified cyclosporin synthetase, to methods for the production of cyclosporin-like peptides or derivatives using the modified cyclosporin synthetase and to the cyclosporin-like peptides or derivatives.

The fungus *Tolypocladium inflatum* GAMS (also known as *Tolypocladium niveum*) produces cyclosporins, a group of neutral cyclic peptides composed of eleven amino acids. Other fungi have been found which may form cyclosporins but *Tolypocladium inflatum* is the most important organism for the production of cyclosporins by fermentation. Cyclosporins exhibit remarkable biological effects: for example cyclosporin A, the main metabolite, is a potent immunosuppressant (**Borel J**, et al (1976) Biological effects of cyclosporin A: a new antilymphotic agent Agents Actions 6: 468). An enzyme has been identified which catalyses the entire peptide biosynthesis of cyclosporin and is therefore called cyclosporin synthetase (**Zocher R**, et al (1986). Biosynthesis of Cyclosporin A: Partial Purification and Properties of a Multifunctional Enzyme from Tolypocladium inflatum. *Biochemistry* **25**, 550-553 and **Billich** A et al (1987) Enzymatic Synthesis of Cyclosporin A J. Biol. Chem. 262, 36, 17258-17259). The biosynthesis proceeds non-ribosomally by a thiotemplate process, as has also been described for other peptide synthetases (**Kleinkauf H** et al (1990) Nonribosomal biosynthesis of peptide antibiotics Eur. J. Biochem. 192, 1-15). Each amino acid is first activated in the form of an adenylate, then bound in the form of a thioester and linked with the following amino acid to the peptide. In the case of cyclosporin A, seven of the amino acids, bound as thioesters, are methylated before they are linked to the preceding amino acid in a peptide bond. This methylation function is an integral constituent of the enzyme polypeptide (Lawen A et al (1990) J. Biol. Chem. 265, 11355-11360). Including the cyclisation reaction, cyclosporin synthetase performs at least 40 reactions.

Cyclosporin A contains three non-proteinogenic amino acids: D-alanine in position 8, a-amino butyric acid in position 2 and, in position 1, the unusual amino acid (4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine (Bmt or C9 amino acid). All three amino acids must be each prepared by a biosynthetic pathway which is independent of the primary biosynthetic pathway. Cyclosporin synthetase does not possess an alanine-racemase function (**Kleinkauf H** et al (1990), as above) and thus, D-alanine cannot be produced by cyclosporin synthetase by

CONFIRMATION COPY

epimerisation of enzyme-bound L-alanine, as is the case for other peptide antibiotics whose biosynthesis mechanism is known.

European patent application number having publication number EP 0,578,616 discloses an isolated DNA sequence which codes for an enzyme having cyclosporin synthetase-like activity. The whole contents of this application are included herein by reference thereto.

According to the present invention, as hereinafter described, amino acid sequence comparisons between domains of cyclosporin synthetase are used to identify regions of the sequence which are responsible for the recognition and binding of the individual amino acids. Using a limited set of selection rules it is possible to identify three amino acid positions in the subdomain sequences which are responsible for amino acid specificity. Homology with the firefly luciferase protein shows that these three key residues are close to each other and line the surface of a putative specific substrate binding pocket located on the amino acyl-adenylation subdomain. The method of the present invention allows the prediction of a large number of cyclosporin synthetase mutants. These cyclosporin synthetase mutants enable the synthesis of cyclosporin-like peptide. Cyclosporin synthetase (CySyn) is the most complex of a constantly rising number of peptide synthetases known so far (**Stachelhaus T.** et al (1995) FEMS Microbiol. Lett. 125, 3-14). It is a single chain polypeptide consisting of 15281 amino acids with a deduced molecular mass of 1.69 Mda (**Weber G** et al (1994) Curr. Genetics 26, 120-125). The protein is produced by the fungus *Tolypocladium inflatum Gams* for the non-ribosomal synthesis of the undecapeptide cyclosporin A (CsA) (Figure 1) (**Dittmann J.** et al (1994) J. Biol. Chem. 269, 2841-2846). CsA is used clinically as an immunosuppressant to prevent organ rejection following transplant operations (**Borel, J.** et al (1989) Pharmacol. Rev. 41, 239-242).

The sequence of CySyn shows a repeating structure with 11 homologous modules responsible for the incorporation of each of the eleven amino acids in the product (Figure 1) and a C-terminal "twelfth domain" putatively used for the cyclisation reaction of CsA. A "module" contains the domains (e.g. adenylating, thiolation and optionally methylating domains), corresponding to a particular amino acid. A domain is part of a module.

The bacterial antibiotics including gramicidin S, tyrocidine, surfactin and bacitracin, and the fungal peptides HC-toxin, enniatin A and cyclosporin A are all synthesised by multifunctional enzymes which have been characterised and reviewed (**Stachelhaus, T.** et al (1995), as above, **Kleinkauf, H.** et al (1996) Eur. J. Biochem. 236, 335-351)). Comparison of the amino acid sequences of the different synthetases defined a multidomain architecture

(**Weber G** et al (1994), as above). The eleven large modules of CySyn are responsible for the adenylation, thioesterification and (for seven of the amino acids) N-methylation, of each amino acid in the growing chain. It has been shown that the first module in CySyn is responsible for the synthesis of D-Ala 8 (Figure 1) and that the peptide product is synthesised in a sequential and stepwise fashion (**Dittmann, J.** et al, as above).

A comparison of bacterial and fungal synthetases have led to the identification of core modules [1], which have some degree of evolutionary conservation (**Cosmina, P.** et al (1993) Mol. Microbiol. 8, 821-831). Most important was the finding, that each domain within the different modules of various synthetases can act independently and can perform consecutive sequential steps like multienzyme complexes (**Stachelhaus, T.** et al. (1996) Chemistry and Biology 3, 913-921). Prior to the present invention, regions corresponding to substrate binding pockets for the synthetases have however not previously been identified.

The present invention provides a method of modifying a domain of cyclosporin synthetase to generate a cyclosporin synthetase having an altered amino acid recognition specificity comprising modifying at least three amino acids in the domain.

In a further embodiment of the present invention there is provided a method of modifying the amino acid recognition specificity of cyclosporin synthetase comprising site directed mutagenesis of the three residues a, b and c as shown in Figure 3, in at least one of the domains 1 to 11 shown in Figure 3.

In a preferred embodiment of the method of the present invention, the amino acid specificity is selected from the following amino acid specificities: Leu, Val, Bmt, D-Ala, L-Ala, Abu and Gly.

In a further embodiment of the present invention there is provided a method of modifying cyclosporin synthetase to generate a modified cyclosporin synthetase having an altered amino acid recognition specificity comprising changing an amino acid at at least one position of the three residues a, b and c as shown in Figure 3 to a different amino acid, in at least one of the domains 1 to 11 shown in Figure 3.

In a yet further embodiment of the present invention there is provided a method of making a modified cyclosporin synthetase, comprising inserting a site directed mutation at at least one of the three residues a, b and c as shown in Figure 3, in at least one of the domains 1 to 11 shown in Figure 3, said synthetase being capable of generating a cyclosporin-like peptide

In a yet further embodiment of the present invention there is provided a method of modifying cyclosporin synthetase wherein the specificity of the D-Ala domain is altered to L-Valine. In a preferred embodiment the modified cyclosporin synthetase has a "bc" mutation.

In a yet further embodiment of the present invention there is provided cyclosporin synthetase obtainable according to a method of the present invention.

In a yet further embodiment of the present invention there is provided a method of generating a cyclosporin-like peptide comprising making a modified cyclosporin synthetase and expressing said synthetase in *Tolypocladium inflatum* and isolating the cyclosporin-like peptide from *Tolypocladium inflatum* culture. Peptides produced according to this method are also part of the present invention, for example a cyclosporin-like peptide wherein the D-Ala of cyclosporin-A is replaced by L-Val.

The present invention further provides cyclosporin-like peptides made by the cyclosporin synthetase of the present invention.

The identification of the site-specific amino acid binding pockets by dissecting the cyclosporin synthetase protein into its various domains followed by cross-comparing these subdomains with each other is described hereinafter. The identification of the site-specific amino acid binding pockets enables the modification of CySyn to generate a modified CySyn which retains the CySyn function, namely the ability to make CsA-like peptides. By the term "modification" of CySyn is meant that the amino acid at at least one position of CySyn is changed to a different amino acid. The amino acid modification can be effected by any conventional method of modification known to one skilled in the art e.g. site-directed mutagenesis. It will be clear from the above to one skilled in the art, that according to the present invention the CySyn can be modified such that one, or more amino acids can be changed to generate a modified CySyn capable of making Cs-A-like peptides.

Homology Searches and Protein Alignment: Initial CySyn subdomain homology searches are done using a BLASTP search engine (**Altschul S** et al. (1996) Meth. Enzymol. 266, 460-480) on a non-redundant protein sequence database (http://www.genome.ad.jp). Protein sequences of acyl-adenylation proteins and acyl carrier proteins are extracted from the GENPEPT database (http://www.genome.ad.jp) using the program MPSRCH (**Coulson** et al (1987) Comp. J. 30, 420-424). These sequences are used as an input to the PHD program (**Rost B** et al. (1994), as above) and used for the secondary structure prediction. The program SSS_align (**Sturrock S** (1997) PhD Thesis, Edinburgh University). incorporates the secondary structural information with sequence information to give a good alignment of the ACP

subdomains and the amino acyl- adenylation subdomains with the secondary structures of their respective template proteins.

Molecular Modelling: The molecular graphics program WITNOTP may be used to generate pictures of the luciferase structure showing the location of the amino acids responsible for providing specificity of amino acid substrate binding to the amino acyl adenylation domain.

The internal homology between the eleven repeating units of the CySyn sequence is determined using the program ALIGN (**Collins, J.** et al (1990) Meth. Enzymol. 183, 474-487) and initial boundaries between the different CsA synthetase domain modules are outlined. In order to predict the secondary structure effectively, the following procedure is used. A search of the GENPEPT database is made using the MPSRCH program (**Coulson** et al, as above) and 30 similar sequences for each of the modules are retrieved. These are transferred and used as input into the PHD prediction program ((**Rost, B.**, et al. (1994) Comp. Appl. Biosci. 10, 53-60, **Rost, B** et al (1993) J. Mol. Biol. 232, 584-599)). The strength of this prediction algorithm is in having a wide range of related sequences which makes the boundary between regions of secondary structure more apparent. The program SSS_align (**Sturrock S,** as above) is then used to line up the sequence of the synthetase domains using both amino acid sequence information and secondary structure information. Comparison of the individual subdomains then allowed the dissection of the whole CySyn molecule into its modules and domains and the definition of clear boundaries between the different modules. The results of this alignment are shown in Figure 2 and summarised in Table 1. Amino acid identity between the domains is high and ranges between 50 and 54%.

The modular structure of CySyn can be divided essentially into two groups (Figure 2). Group I modules (CySyn modules 1, 6, 9 and 11 which code for CsA amino acids 8, 2, 5 and 7 respectively) and the other modules where the corresponding positions in the final drug product are N-methylated. All CySyn modules start with an N-terminal domain (NTD) of about 460 amino acids, which may be involved in the elongation mechanism (condensation reaction) (**Decrecylagard, V.** et al (1995) Life Sciences 318, 927-936) Database searches revealed no homology with other known proteins. Module 1 has a small 20 amino acid extension at its N-terminus. There is also a C-terminal domain at the end of module 11 which shows no homology with other domains.

The acyl-adenylation domain consists of about 450-550 amino acids and has significant homology with all peptide synthetases (**Stachelhaus T.** et al (1995), as above) as well as CoA

synthetases and other adenylating enzymes including firefly luciferase (**Conti, E** et al (1996) Acta Crystallog. D 52, 876-878). Homology between these ATP-binding domains is between 20% and 80% amino acid identity. There are eleven acyl-adenylation domains in cyclosporin synthetase and each is responsible for recognition of the cognate amino acid substrate and the formation of an activated amino acyl-adenylate by reaction of the amino acid carboxy group with bound ATP.

Seven of the eleven amino acids in CsA are N-methylated (Figure 1) and each of the seven relevant CySyn modules have an additional N-methylation domain of about 400-450 amino acids at the C-terminus of the acyl-adenylation domain (**Weber G** et al (1994), as above). The other four modules have a short peptide linker of 13 to 37 amino acids.

In all eleven modules these domain stretches are followed by a synthetase specific peptide with a length of approximately 54 amino acids, which could be found only in other peptide synthetases both from bacterial and fungal origin. The homology of this subdomain within the various modules varies between 50 to 70%, and the identity to other synthetases is approximately 50%. The biological function of this domain is not known.

All eleven CySyn modules have an Acyl Carrier Protein (ACP) domain of 75 amino acids. ACP proteins are also involved in the synthesis of both polyketides and fatty acids (**Simpson, T.** J. (1995) Polyketide biosynthesis. Chemistry and Industry 407-41), and three dimensional NMR structures have been elucidated for actinorhodin polyketide synthase ACP (**Crump, M.** et al (1996) FEBS Lett. 391, 302-306) and *E. coli* fatty acid synthase ACP (**Holak, T.** et al (1988) Eur. J. Biochem. 175, 9-15).

Amino Acid Specificity Pocket. Invariant residues involved in ATP binding among a wide range of acyl adenylation proteins have been identified (**Pavelavrancic, M.** et al (1994) J. Biol. Chem. 269, 14962-14966) and could also be found in all 11 CySyn domains. The cyclosporin A sequence has four occurrences of MeLeu, one occurrence of valine and one of MeVal. The mechanism for peptide synthesis is known to involve the recognition of the unmodified amino acid (**Lawen, A.** (1990) as above) which reacts with ATP to form an acyl-adenylate with the release of pyrophosphate. Such a P-O bond is formed between the amino acid and the AMP at each of the eleven modules. The specificity of the reaction is likely to depend on the shape of the surrounding enzymatic pocket of the amino acyl-adenylation domain. This suggests that the four MeLeu domains should share identical specificity pockets

which are different from all the pockets of the other domains. The sequences of the eleven aligned domains are searched for occurrences of amino acids combinations unique to the four MeLeu domains or unique to the two valine domains. In addition to CsA, the fungus *T. inflatum* produces many other minor metabolites of the same structural type (**Wenger, R.** (1990) Transplant. Proc. 22, 1104-1108). Another constraint therefore included the cross-comparison of the other naturally occurring cyclosporins (i.e. CsB to CsZ). As shown in Table 1, the only amino acid used at position 3 of cyclosporin is glycine and a D-configured residue at position 8. Other residues in the drug can vary to some extent in the incorporated amino acid (**Wenger, R.,** as above). The aligned sequences are searched for occurrences of the following conditions:

1) An amino acid at a given position in the aligned sequence of domains 1 (D-Ala 8) or 7 (Sar 3; Sar is sarcosyl/methylglycine) is different from amino acids in all other domains at the same aligned position.

2) Amino acids at a given position in the aligned sequences of domains 2 (MeLeu 9), 3 (MeLeu 10), 8 (MeLeu 4) and 10 (MeLeu 6) are identical to each other and different from amino acids in all other domains at the same aligned position.

3) Amino acids at a given position in the aligned sequences of domains 4 (MeVal 11) and 9 (Val 5) are identical to each other and different from amino acids in all other domains at the same aligned position.

Positions in the aligned CySyn sequence which meet these criteria are likely to play an important role in the recognition and discrimination of amino acid binding to the amino acyl-adenylation domains.

Only three such amino acid positions are found in the complete sequence of the synthetase modules (Figure 3). They are all found in the amino acyl-adenylation domain located near the ATP binding site. This strongly implies that these residues are crucially important for the recognition and binding of the amino acid substrate. The example below demonstrates that modification of at least one of these residues may result in a change of activity and/or specificity.

Firefly luciferase homology with the aminoacyladenylation domain. The x-ray crystallographic structure of firefly luciferase is known (**Conti, E** et al (1996), as above). This protein is known to have homology with the acyl-adenylation domain of peptide synthetases. The program SSS_align (**Sturrock S,** as above) is used to line up the sequence of the

synthetase domain and the luciferase using both amino acid sequence information and secondary structure information. The results of this alignment are included in Figure 3. The three amino acid positions a, b and c mentioned above are then mapped on the three dimensional model of luciferase. The three residues form an almost perfect equilateral triangle with the Cα atoms separated by between 6.1 Å and 6.9 Å. A stereo picture of these three residues in the X-ray structure of luciferase is shown in Figure 4a.

The three amino acids co-defining the putative amino acid specificity pockets in each of the of CySyn domains are given in Table 2. These residues (labelled a, b, and c in Figure 3) are derived from the simple selection rules described above. They provide a recognition code ("recognition triplet") for the amino acid substrates of the different CySyn amino acyl-adenylation domains. The recognition triplet lining the Me-Leu recognition pocket consists of alanine (A), glycine (G) and methionine (M). These residues have been modelled into the X-ray structure of luciferase and a stereo picture is given in Figure 4b. It is significant that the size of the recognition triplet is inversely proportional to the size of the cognate substrate side chain. Thus the relatively bulky leucine substrate with five non-hydrogen atoms in the side chain fits into a relatively spacious pocket with the recognition triplet contributing only five non-hydrogen side-chain atoms (one from alanine and four from methionine). In contrast when the side chain is glycine, the recognition triplet (IVQ) is rather bulky and a total of 13 side chain atoms fill the recognition pocket. Figure 4c shows the recognition pocket for sarcosine with F247I, G316V and S347Q modelled such that the side-chains adopt their most commonly observed conformations.

The triplet code was originally derived without any knowledge of the possible 3D structure of the amino acyl-adenylation domains. That the positions of the three residues lie in such a clearly defined pocket near the ATP binding site of luciferase provides significant support for the original hypothesis. The recognition hypothesis provides new possibilities for designing cyclosporin-like products. Relatively few site-point mutations in the synthetase gene should enable the production of designed sequences in which, for example, specific leucine residues in cyclosporin could be changed to valine by a double mutation of G to A and M to L as outlined in Table 2. There is also scope for designing new amino acid specificities by using site-directed mutagenesis to engineer new triplet combinations into the recognition pocket of CySyn.

Table 1

Modular structure of cyclosporin synthetase and the substrate specificity of its activating units.

Table 2

Amino acids involved in the specificity pockets of the various cyclosporin synthetase domains. The three sites in the AAA domain (a, b and c) obeying the three substrate specificity rules (see text) have a triangular location to each other (as modelled into the luciferase structure, see Figure 4) creating a putative active site. The corresponding amino acids found in the luciferase molecule site are also included.

Figure 1

Cyclosporin A, a natural undecapeptide from *Tolypocladium inflatum.*

Figure 2

Schematic diagram of the domain organisation of individual cyclosporin synthetase modules. The numbers to the left indicate the modules according to their arrangement in CySyn. Modules 1, 6, 9 and 11 all share the same sequential order including an N-terminal domain (NTD), the amino acyl-adenylation domain (AAA), a short spacer (S) followed by a synthetase specific domain (SSD) and the acyl carrier protein homology region (ACP). In addition module 1 has an N-terminal extension (NT) and module 11 has an extra C-terminal domain (CTD). Modules 2, 3, 4, 5, 7, 8 and 10 have instead of the spacer an N-methyltransferase domain (NMT). Both N and C termini are depicted by N and C respectively.

Figure 3

Multiple amino acid sequence alignments of all 11 amino acyladenylation domains of cyclosporin synthetase and firefly luciferase. A site which was found to be homologous in other synthetases (core 2) and the AMP binding motif (AMP site) are both highlighted. Identical amino acids in all 11 CySyn domains are indicated in CySyn domains 2 to 11 by a period. Gaps and extended loops are introduced in the sequences for maximum alignments and are shown by a hyphen. The three sites which obey the amino acid specificity rules in order to find an active site are boxed and labelled a, b and c. The secondary structure prediction of CySyn domain 11 and the elucidated secondary structure of luciferase are included using 'h' for alpha helices and

'e' for extended conformations. The amino acid positions are depicted on either side of the sequences.

Figure 4

a) Stereo picture of a detail of the X-ray structure of firefly luciferase drawn using MOLSCRIPT (**Kraulis, P.** (1991) J. Appl. Crystallog. 24, 946-950) and showing the positions of the three residues defined to be most important in substrate side-chain recognition in the homologous domains of cyclosporin synthetase.

b) As in 4a) but showing the 'recognition triplet' for a cyclosporin leucine side-chain. The F247A and S347M mutations of the luciferase X-ray structure are carried out using the modelling program WITNOTP.

c) As in 4a) but showing the 'recognition triplet' for sarcosine with the F247I, G316A and S347Q mutations.

Figure 5 A

Restriction map of the Expression vectors pQE60 of Qiagen. In pQE60 the start ATG is part of an *NcoI* recognition sequence. Since an *NcoI* site is also part of cloned D-Ala module, ligation is by blunt ends if pQE60 is used. Accordingly vector pQE60 is cut with *NcoI* and the resulting 5' overhanging ends are completed using a Klenow fragment E. coli DNA-polymerase I . "Leserahmen" means "reading frames".

Figure 5 B

Cloning scheme for the construction of the plasmid pQE-60atgala. The D-Ala module is amplified from genomic DNA using PCR. The primers Dom 8-5 and Dom 8-6 are positioned so that the amplified fragments can be cloned in the correct reading frame after the start ATG. The Dom 8-6 primer is modified at its 5' end by attaching a BamH1 restriction site. "Überhang" means "overhang".

Figure 6

Illustration of the site directed mutagenesis of the three positions a, b and c. The horizontal arrow ATG-D-Ala symbolises the expressed D-Ala module of Cyclosporine synthetase- The vertical blocks show the positions of the core regions 1 to 5 in the adenylation domain and the block marked "core 6" marks the position of the thioester domain (**Marahiel M,** (1992) Multidomain enzymes involved in peptide synthesis. FEBS, Vol. 307,1992 (1), 40-43). The ATP binding site is

marked as "AMP-site". The three amino acids in the substrate binding pocket of D-Ala module are shown from left to right (leu, ile, val) in the shaded boxes. The sequence mutations are indicated in each of the shaded boxes.

Figure 7
Illustration of the specific activity (cpm/μg) of mutations and wild-type protein (D-Ala-ATG)

Figure 8
Illustration of mutation "d".

Figure 9
Specific activity of the "d" Mutations and the Wild type proteins.

Some examples of new triplet combinations are discussed below. Modifications are single, double, triple or quadruple mutations.

Example 1:
Point mutations of D-Ala Module whereby the specificity for D-Ala is designed to change to L-Val

Materials

E.coli M15 (Nal^S, Str^S, Rif^S, lac^-, ara^-, gal^-, mtl^-, F^-;$recA^+$,uvr^+) and SG13009 (Nal^S, Str^S, Rif^S, lac^-, ara^-, gal^-, mtl^-, F^-;$recA^+$,uvr^+), both of which are commercially available from Qiagen, are used for expression of the Cyclosporin synthetase module. *Escherichia coli* XL1 Blue (Stratagene) or DH10b (F^-, *mcr*A, Δ (*mrr-hsd*RMS-*mcr*BC), Φ80d*lac*ZΔM15, Δ*lac*X74, *deo*R, *rec*A1, *end*A1, *ara*D139, Δ(*ara*, *leu*)7697, *gal*U, *gal*K, λ^-, *rps*L, *nup*G) (Life Technologies), routinely used in cloning experiments are used for intermediate cloning.

Conventional techniques for cell culturing, cell disruption (e.g. ultrasound), DNA manipulation, isolation of plasmids and DNA, elution of DNA fragments from agar gels, electrotransformation(e.g. gene pulsing), production of electrcompetent E-coli cells, dephosporylation of DNA after restriction by restriction enzymes and purification of the

dephosphorylated DNA, ligation (e.g classic techniques using T4 ligase as described by **Maniatis** et al (Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor 1982.) using for example the Ligation Express™ Kit of Clontech, polyacrlamide gel electrophoresis, sequencing of DNA, hybridisation, clone hybridisation and clone isolation, sequencing and analysis of DNA sequences are employed.

Different *Taq*-Polymerases (Boehringer, Promega) are used according to the manufacturer's instructions for analytical PCR. For production of plasmid DNAs, LONG EXPAND Polymerase MIX (Boehringer) and buffer Nr.1 (Boehringer) are used. The conditions for the PCR-Programs for the primer combinations G05+G06, G07+G08 and G09+G10 are: 1. 95°C 2 min, 2. 95°C 10 sec, 3. 60°C 1 min. 30sec, 4. 72°C 8 min, 5. 72°C 5 min, 6. 4°C on-hold. Steps 2, 3 and 4 are repeated 35 times. To purify the resultant PCR products, 1 ml purification resin (Magic, Promega) is used and passed over a mini-column, washed with wash solution and finally the PCR product is washed again from the column.

The D-Ala activating module of CySyn is modified. The module is purified and renatured using inclusion bodies (IB) as described by **Rudolph R** et al. (1996) "Protein Funtion: A Practical Approach", Ed. Creighton, T.E. 1-39 (in particular see working example number 9 "Standard Protocol for Dissolution and IB Preparations"). The thus obtained IBs are dissolved in 7MGdm/HCl,0.1MTris/HCl,50mMDTT,5 mM EDTA, pH 8.0 and finally purified by passing through a SEC-Chromatographer in 6 M ammonia, 0.1 M Tris/HCl, 50 mM DTT,5 mM EDTA, at pH 8.0. The purity of the fractions is checked by SDS-PAGE.Those fractions containing the protein in homogenous form, are pooled for final renaturation. The optimum conditions for renaturation are 20mM Tris/HCl, 15% (w/v) sorbitol, 5mM EDTA, 2mM DTT, pH 8.0. Chromatography is carried out at 4°C and the fractions are incubated for a further 24H at the same temperature. The activity of the fractions is determined according to the method disclosed by **Martin G** et al: "Analysis of Core Sequences in the D-Phe Activating Domain of the Multifunctional Peptide Synthetase TycA by Site-Directed Mutagenesis". J.Bacteriol. 176 (1994) 2654-2662.

To demonstrate the change in the specificity of the D-ala module from D-Ala to L-Val, special attention is taken to ensure that the new codons are not misread during expression in *Escherichia coli.* A suitable expression plasmid such as the *pQE-* vector system of Quiagen is used to express the D-ala module in *E.coli.* This plasmid has previously been used in

purification of peptide synthetase domains (**Staehelhaus T** et al Chemistry & Biology 3 (1996) 913-921). The "Agtala" fragment corresponds to the first module of CySyn and is cloned from the postulated start codon of CySyn (**Weber G** et al (1994), as above). The peptide synthetase catalysing cyclosporin production in *Tolypocladium niveum* is encoded by a giant 45.8 kilobase open reading frame. (see **Weber** et al (1994), as above). The end of the fragment is as described by the relevant sections of **Knes**, O.: "Klonierung einer Domäne aus dem Cyclosporinsynthetasegen aus Tolypocladium niveum und ihre Expression in Escherichia coli." Diplomarbeit, Technische Universität Graz, 1996). An illustration of the plasmid is given in Figure 5A (for more details see relevant sections of Rützler M.: "Klonierung und Charakterisierung der D-Alanin aktivierenden Einheit der Cyclosporinsynthetase aus *Tolypocladium inflatum.*" Diplomarbeit, Universität Salzburg 1997.). Figure 5B illustrates the cloning scheme for the construction of pQE-60atgala. For a more detailed description of the construction of this plasmid please refer also to the relevant sections of Rützler M.: Klonierung und Charakterisierung der D-Alanin aktivierenden Einheit der Cyclosporinsynthetase aus *Tolypocladium inflatum.* Diplomarbeit, Universität Salzburg 1997).
E.coli M15 which contains the plasmid pREP4 and comprises the gene for the *lac* repressor, *lacI* , is used.

Figure 6 illustrates the site directed mutagenesis of the three positions a, b and c.
PCRs with the following primer combinations G05 + G06, G07 + G08 and G09 + G10 are carried out with plasmid DNA of the expression construct pQE-60atgala. PCR products of about 6.7 kb are amplified. The purified PCR fragments are cut with the relevant restriction enzymes (G05 + G06 with *Mlu*I, G07+G08 with *Ksp*I and G09+G10 with *Bfr*I). Following purification by gel electrophoresis, the isolated bands are religated and introduced into a suitable *E.coli* host e.g. *E. coli* DH10b or *E. coli* XL1 Blue. The resulting clones are pQE-60atgala-a" or "clone a" (G05 + G06), "pQE-60atgala-b" or "clone b"(G07 + G08) and "pQE-60atgala-c" or "clone c" (G09 + G10)." Starting from "clone a", PCRs with G07 + G08 and G09 + G10 are carried out in order to obtain clone "ab" and "ac" respectively. Single DNAs are verified by partial sequencing Plasmid DNA, the expression of which is checked, serves as the starting product for introducing further point mutations. For example starting with clone "a", PCRs with the primer combinations G07 + G08 and G09 + G10 are carried out in order to obtain clones "ab" and "ac" respectively. Thereafter, PCR is carried out on clone "ab" to obtain clone "abc". In this way, each of the seven clones ("a", "b", "c", "ab", "ac", "bc" and "abc") is obtained. The activity and specificity

of the mutations is analysed using radioactive ATP/PPi exchange experiments as above mentioned. As is shown in Figure 7, the wild-type D-Ala module possesses a pronounced specificity for D-Ala and shows practically no activity for L- or D-valine. Mutations "b". "ab" and "abc" are practically inactive. Mutations "a" and "ac" show reduced activity in respect of D-Ala and no valine activation. Mutation "c" activates L-valine to a very slight extent whilst still showing very high activation of D-ala. Mutant "bc" activates L-valine to significantly better than D-Ala. The D-Ala value is still very clear (some 1/7 of the activity of the non-mutated variant) but the specific activity towards L-valine is some 30 times higher. Mutant "bc" recognises and activates D-valine to a slight extent. It is clear, that the above mutations, in particular mutant "bc" are involved in the recognition and specificity thereof ,of the substrate amino acids. It should be appreciated that the value in relation to specific and activating activity as shown in Figures 7 and 9, should be interpreted as an indication of a tendency.

Example 2:

As in Example 1 above, combinations of a further site-directed mutation are carried out, the strategy of which is shown in Figure 8, namely mutation "d". The production of the mutations illustrated in Figures 8 and 9, is carried out using primers G11 and G12 (see Table 3) which on the one hand at the protein level comprise the mutation of Phe to Met and on the other hand include a single restriction site for the restriction enzyme *Nru*I. Starting with pQE-60atgala, and using the primer combination G11 and G12, PCR are carried out under the following conditions: Step 1. 95°C 2min, step 2 . 95°C 10sec, step 3. 58°C 30sec, step 4. 68°C 6min, step 5. 95°C 10sec, step 6. 58°C 30sec, step 7. 68°C (starting for 6min. 68°C, + 20sec. each time this step is repeated), step 8. 68°C 15min, step 9. 4°C on-hold. Steps 2, 3 and 4 are repeated 10 times, steps 5, 6 and 7 are repeated 30 times. The resulting PCR products are purified as above and cut with the restriction enzyme *Nru*I. The remaining steps are carried out as described in Example 1. In this way, all of the mutations shown in Figure 9 apart from "bd", are obtained. Mutant "bd" is obtained by starting with the verified mutation "d" clone and introducing mutation "b" using primer combination G07 and G08 and the PCR conditions given above for this primer combination. As is seen from Figure 9, mutations "d" and "bd" are practically inactive, mutations "cd" and "abd" show a clearly reduced activity for D-Ala, mutations "ad" and "cad" show significant activation for D-valine, but not for L-valine, mutations "cbd" , "abcd" and "bc" possess the highest activity towards L-valine.

Whereas in the case of "bc" the additional mutation "d" results in a worsening of activity, in the case of "abc" the domain is changed from a inactive form (see Figure 7) to an active form. From the above results it is clear that position "d" is also involved in substrate specificity.

Example 3: Production of cyclosporin-like peptide.

A modified D-Ala module obtained as set out in Example 1 above is exchanged for the wild type D-Ala module in CySyn. The thus modified CySyn is introduced into *Tolypocladium inflatum* using methods known to those skilled in the art, preferably using a method for gene replacement. The *Tolypocladium inflatum* is incubated using conventional materials and methods. After a suitable period, a cyclosporin-like peptide, in which the D-Ala of Cyclosporin-A is replaced by L-Valine, is isolated from the *Tolypocladium inflatum* culture.

As outlined in Example 3, it is possible to exchange a wild type domain e.g. a D-Ala domain for a modified D-Ala domain. It will be understood, that in an analogous way it is also possible to exchange another domain/module or also more than one wild type domain/module. Exchange of more than one domain/module results in a cyclosporin synthetase having a corresponding number of altered amino acid recognition specificities. In this way, it is possible to arrive at a number of different cyclosporin-like peptides. It will be understood that the applicant does not wish to be limited by the above examples.

Tables and Figures

Table 1

Module[1]	CsA aa[2]	Cs(n) aa[3]	length[4]	NTD[5]	AAA[6]	S[7]	NMT[8]	SSD[9]	ACP[10]
1	D-Ala	D-Ser	1104	21-481	482-935	936-973		974-1028	1029-1104
2	MeLeu	Leu	1498	1105-1567	1568-2022		2023-2471	2472-2526	2527-2602
3	MeLeu	Leu	1487	2603-3064	3065-3515		3516-3958	3959-4013	4014-4089
4	MeVal	Val	1492	4090-4550	4551-5005		5006-5450	5451-5505	5506-5581
5	MeBmt	MeLeu	1497	5582-6043	6044-6497		6498-6948	6948-7002	7003-7078
6	Abu	Ala,Thr, Val,Nva	1060	7079-7540	7541-7993	7994-8007		8008-8062	8063-8138
7	Sar		1495	8139-8601	8602-9055		9056-9502	9503-9557	9558-9633
8	MeLeu	Val	1497	9634-10095	10096-10546		10547-10999	11000-11054	11055-11130
9	Val	Nva	1072	11131-11584	11585-12036	12037-12071		12072-12126	12127-12202
10	MeLeu	Leu	1496	12203-12664	12665-13112		13113-13566	13567-13622	13623-13698
11	L-Ala	Abu	1075	13699-14162	14163-14615	14616-14641		14642-14697	14698-14773

[1] Module: Cyclosporin synthetase module number; [2]CsA aa: Amino acid incorporated in CsA at the given position; [3]Cs(n) aa: Amino acid which can be found in other minor metabolites of cyclosporin at the given positions; [4]length: total peptide length of the corresponding synthetase module; [5]NTD: N terminal domain, its start and end positions within the whole protein are indicated; [6]AAA: amino acyl-adenylation domain; [7]S: spacer peptide; [8]NMT: N-methyl transferase domain; [9]SSD: synthetase specific domain; [10]ACP: acyl carrier protein homology domain

Table 2

	a	b	c
Leu	A	G	M
Val	A	A	L
Bmt	A	G	I
D-Ala	L	I	V
L-Ala	V	A	L
Gly	I	V	Q
Abu	A	A	Y
Luciferase	F	G	S

Table 3 Oligonucleotide Primers for PCR

Oligo (Sequence No.)	*Length*	*Sequence in the 5' to 3' direction*
Dom 8-5 (1)	30mer	ggc gcc atc ggg caa gac atg gca tat gat
Dom 8-6 (2)	35mer	ggg gat ccc atc ggc tgc tca aag aca gtt gct ac
G05 (3)	34mer	g ggg gac gcg tca aca tgg gag atc tat acc cct
G06 (4)	34mer	g ggg gac gcg tcg aaa gcg atg ttg gcc agg tga
G07 (5)	34mer	g ggg gcc gcg ggc gat cgc ttc gac cga cgt gac
G08 (6)	34mer	g ggg ccc gcg gcg tac agg cta tcc agg ccc gca
G09 (7)	34mer	g ggg tct taa gta cca tct aca gcg tct ccg agg
G10 (8)	34mer	g ggg act taa gac gga att ctc ggt tgg acc ata
G11 (9)	40mer	ggg ggt cgc gat gct tgc tcc tgc tct gat caa gca gtg t
G12 (10)	40mer	ggg ggt cgc gac acg aat gcc ctc ctt ggt gaa tac agc t

CLAIMS

1- A method of modifying a domain of cyclosporin synthetase to generate a cyclosporin synthetase having an altered amino acid recognition specificity comprising modifying at least three amino acids in the domain.

2. A method of modifying the amino acid recognition specificity of cyclosporin synthetase comprising site directed mutagenesis of the three residues a, b and c as shown in Figure 3, in at least one of the domains 1 to 11 shown in Figure 3.

3. A method as claimed in either claim 1 or 2 wherein the amino acid specificity is selected from the following amino acid specificities: Leu, Val, Bmt, D-Ala, L-Ala, Abu and Gly.

4. Cyclosporin synthetase obtainable according to a method as claimed in any one of claims 1 to 3 above.

5. A method of modifying cyclosporin synthetase to generate a modified cyclosporin synthetase having an altered amino acid recognition specificity comprising changing an amino acid at at least one position of the three residues a, b and c as shown in Figure 3 to a different amino acid, in at least one of the domains 1 to 11 shown in Figure 3.

6. A method of making a modified cyclosporin synthetase, comprising inserting a site directed mutation at at least one of the three residues a, b and c as shown in Figure 3, in at least one of the domains 1 to 11 shown in Figure 3, said synthetase being capable of generating a cyclosporin-like peptide.

7. A method as claimed in any one of claims 1, 2, 3, 5 or 6 wherein the specificity of the D-Ala domain is altered to L-Valine.

8. A method as claimed in claim 7 wherein the modified cyclosporin synthetase has a "bc" mutation.

9. Cyclosporine synthetase obtainable according to any one of claims 1, 2, 3, 5, 6, 7 or 8.

10. A method of generating a cyclosporin-like peptide comprising making a modified cyclosporin synthetase and expressing said synthetase in *Tolypocladium inflatum* and isolating the cyclosporin-like peptide from *Tolypocladium inflatum* culture.

11. A cyclosporin-like peptide obtainable according to the method as claimed in claim 10.

12. A cyclosporin-like peptide as claimed in claim 11 wherein the D-Ala of cyclosporin-A is replaced by L-Val.

Figure 1

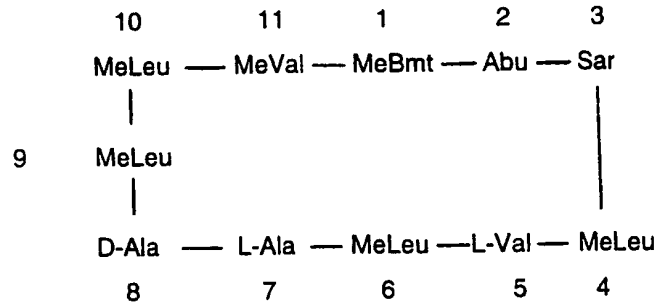


Figure 2

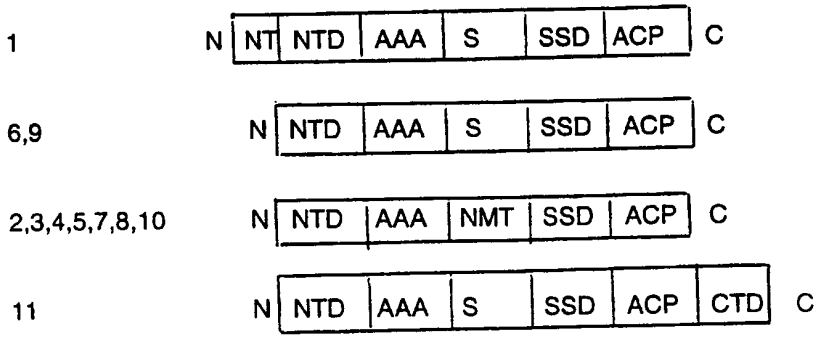


Figure 3

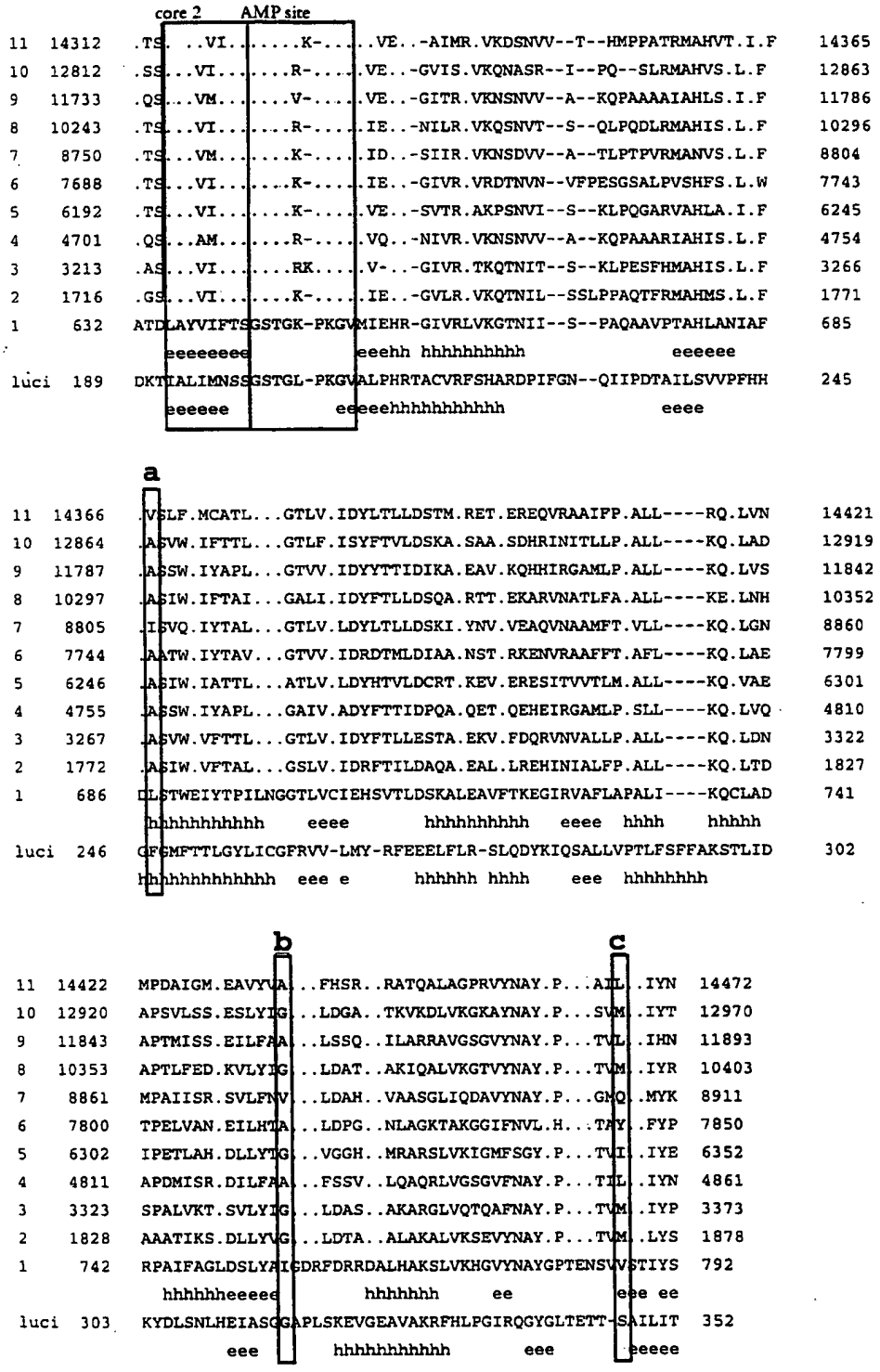
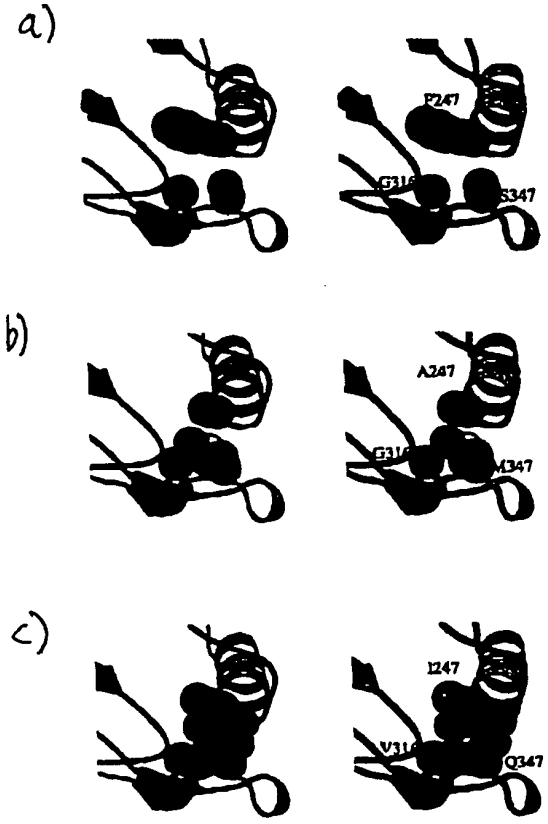
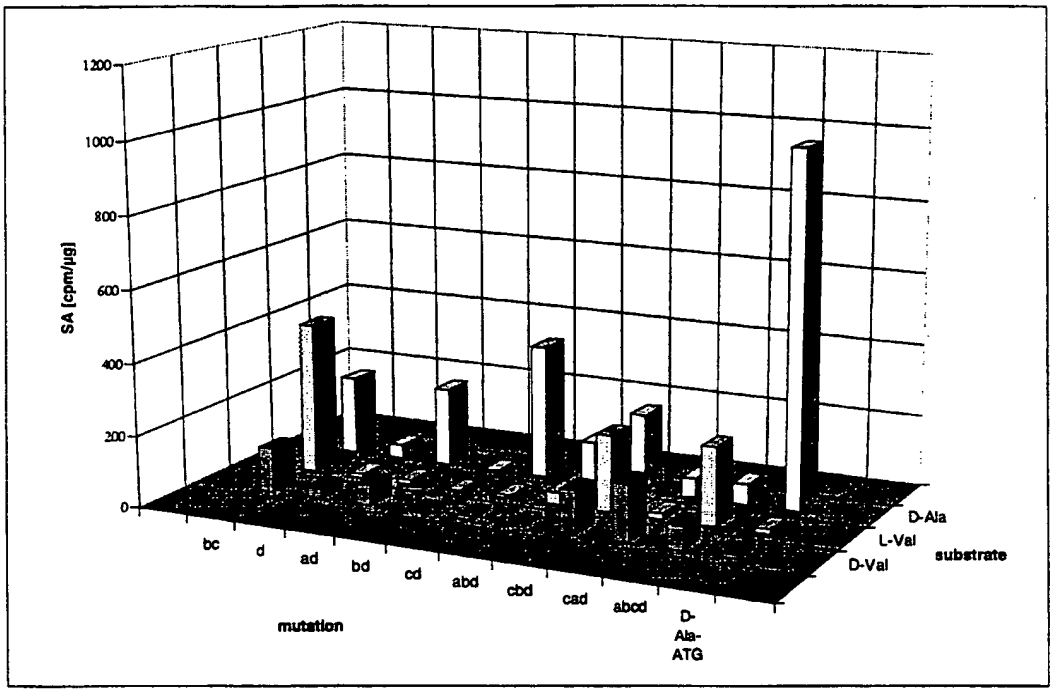


Figure 4



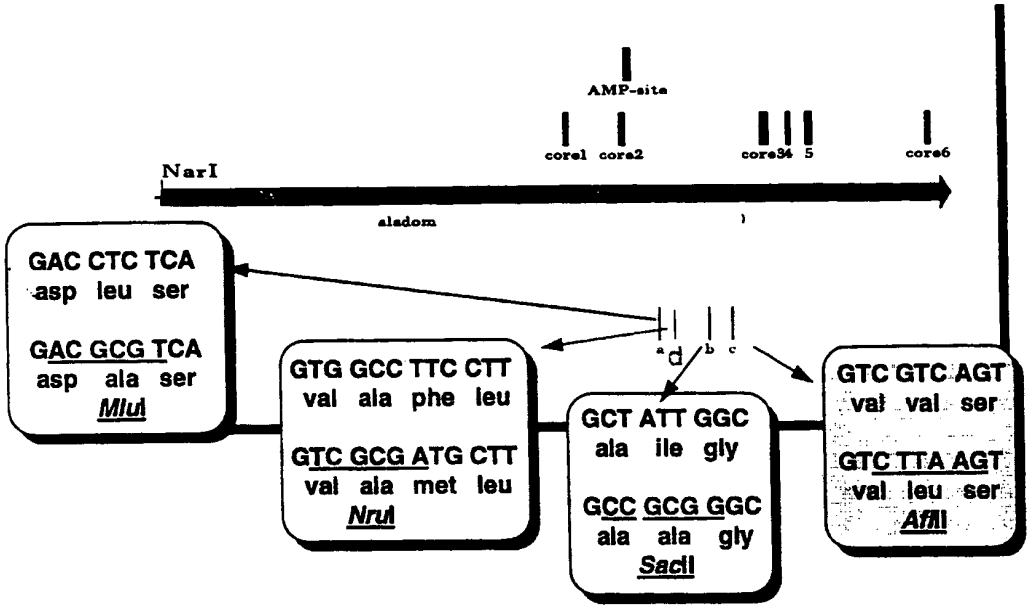
BEST AVAILABLE COPY

Figure 9: Specific activity of the "d" Mutations and the Wildtype proteins.



BEST AVAILABLE COPY

Figure 8 Illustration of Mutation "d"



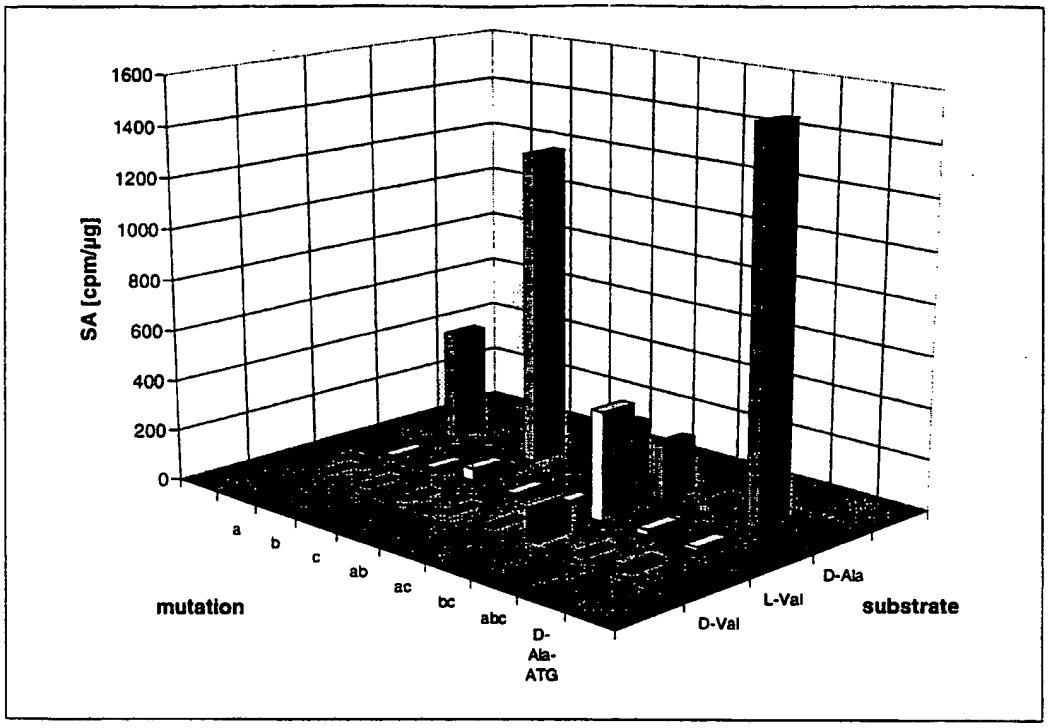


Figure 7

Specific Activity (cpm/µg) of mutations shown and the wild-type protein (D-Ala-ATG)

BEST AVAILABLE COPY

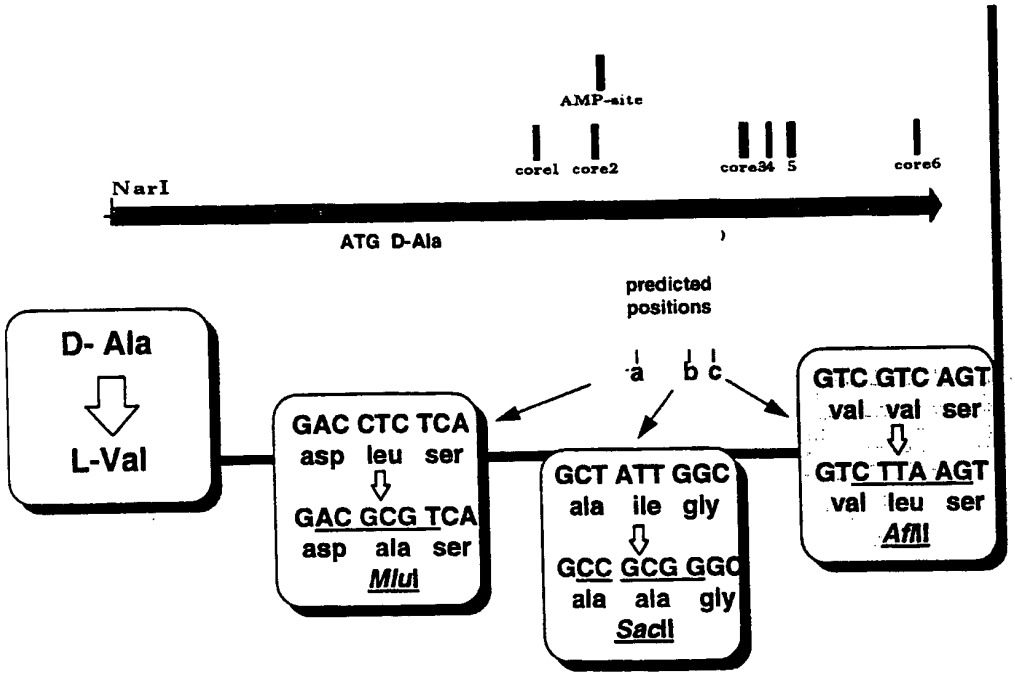
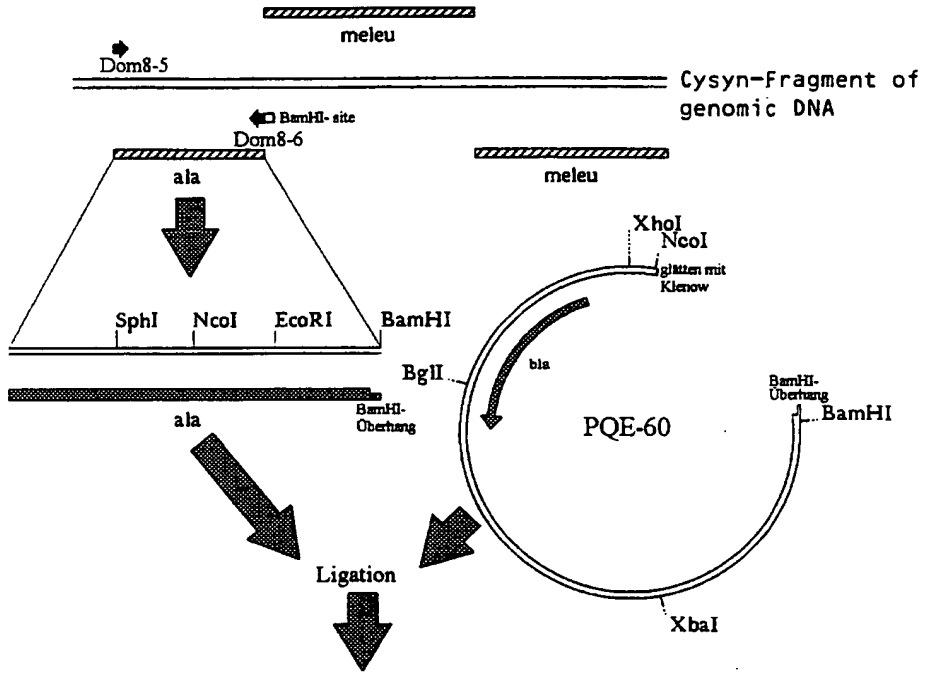


Figure 6 Illustration of site directed mutagenesis of the three positions a, b and c

Figure 5B Cloning scheme for the constrtuction of the plasmid pQE-60atgala



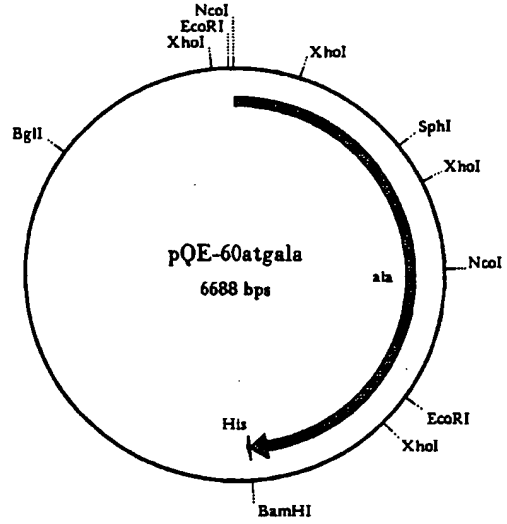


Figure 5A: Restriction map of the Expression vectors pQE60 of Qiagen